Serial ATA

A Promising New Alternative for Enterprise Storage Applications

A JOINT WHITEPAPER PREPARED BY
Eurologic Systems, Intel® Corporation and Western Digital Corporation

Contents

1.0 The Enterprise Storage Challenge

As the demand for data storage explodes, market pressures are mounting on providers of enterprise storage solutions. While enterprise storage capacity requirements are growing at a tremendous rate, doubling every year for the foreseeable future[1], IT budgets are not. In the face of uncertain economic times, IT managers are under intense pressure to find storage solutions that deliver more capacity and density for less money – without sacrificing performance or data availability.

Challenging market conditions are often the catalyst for innovative technology solutions. So it is with the storage marketplace. Coincident with the demand for enterprise storage that meets better price/performance targets, is the emergence of Serial ATA, providing the ability for low-cost ATA devices to grow to higher levels of performance and incorporate enhanced feature sets.

In this paper, we examine the new Serial ATA (SATA) interface and discuss why it provides an attractive alternative for storage systems targeting a range of enterprise storage applications.

1.1 Enterprise Storage Requirements

When designing storage solutions for the enterprise marketplace, system providers must be mindful of the demanding requirements of server and networked storage systems – needs that go beyond those of desktop systems. These critical requirements include:

- **Performance –** Enterprise storage systems must provide high throughput (high data transfer speeds) to serve up large files quickly, maximizing storage system utilization to enhance workgroup productivity.
- **Reliability –** Lost data or lost access to data is a critical concern for any enterprise system, including those handling applications and data not technically considered "mission critical."
- **Availability –** Enterprise storage systems must deliver high uptime and effective failover protection to meet the expectations of storage consumers who expect that their data will always be present and available.
- **Data Integrity –** Data must be accurate and free from corruption.
- **Scalability –** Enterprise storage systems must support the ability to build onto an existing system, in terms of both performance and capacity, while preserving storage investments.
- **Interoperability –** As enterprise IT environments become increasingly heterogeneous, storage system providers must provide solutions that interoperate readily with other storage network elements, systems and applications, using industry standards.

The other critical enterprise priority today is *total cost of ownership* (TCO). Reducing the cost of satisfying the requirements listed above has become an issue of strategic importance-both for storage system vendors and for their customers.

1.2 Classifying Data Types

Of great importance to anyone designing or managing enterprise storage systems is the criticality of the information stored on those systems. As any IT manager knows, there is truly no such thing as "non-critical" data; someone within the enterprise considers all business information of value. While data in the past has been classified according to a range of criticality descriptors, there are today really only two types of information:

- **Mission Critical –** This information is of the highest value and criticality, from both availability and a security viewpoint.

1. The SAN Book 3.0 (Storage Area Networks), Salomon Smith Barney, October 22, 2001.

- **Business Critical –** This represents the great bulk of data stored on enterprise servers and in networked storage systems, consisting of business applications and data accessed continually, and archived data which is accessed only periodically.

The question for storage solution vendors is not how to create low-cost solutions for all enterprise storage applications. Rather, the challenge is how to create more affordable systems with the performance, reliability, scalability, and interoperability required to meet the needs of Business Critical storage applications, which account for the vast majority of enterprise data stores.

2.0 Evolution of Storage Interface Standards

Since its introduction nearly two decades ago, Parallel ATA, also known as IDE (Integrated Drive Electronics), has become the standard storage interface for the majority of personal computers. One reason for its popularity is its relative simplicity, resulting in low-cost storage drives, interface cards and other devices based on ATA technology. Today, nearly all desktop and portable PCs feature IDE disk storage based on the ATA interface standard.

Traditionally, however, ATA did not support the performance, reliability or scalability requirements required by enterprise systems. These storage applications have been dominated by solutions based on the Small Computer System Interface (SCSI) standard, with the more recent entry of Fiber Channel. Systems based on these technologies deliver the high performance and reliability required for mission-critical applications. However, the cost for systems based on these technologies can be substantially higher than those based on ATA.

Due to the cost advantages offered by ATA, solution providers have for some time ventured to produce enterprise storage solutions based on this venerable desktop standard. ATA drives are used commonly in entry-level and blade servers. With a unit CAGR of 69% through 2006,[2] this market segment is forecast to experience the highest growth of all server types (rack, blade, or pedestal).[3] Within the NAS market, numerous vendors have been shipping systems with ATA technology, such as the Quantum Snap!* appliance. In 2002, a number of leading storage vendors have introduced ATA-based subsystems to address the burgeoning tape replacement/near-line storage markets. An example is Network Appliance's NearStore* appliance.

Even server vendors have included ATA options in their entry-level segment including the HP ProLiant,* IBM xSeries* and Dell PowerEdge* product lines.

For the most part, ATA-based systems have been limited to storage applications requiring the lowest performance and reliability requirements – usually steady-state content, such as stored images and archived email. SCSI has continued to dominate the enterprise server and networked storage marketplace.

2.1 Enter Serial ATA

A new specification, called "Serial ATA 1.0," was formally introduced in August of 2001 and is intended as an evolutionary replacement for Parallel ATA. It was first developed as a replacement for desktop ATA. However, storage vendors quickly realized that Serial ATA has distinct advantages over traditional ATA that make it a viable alternative for many enterprise storage applications, enabling storage systems at what current pricing models estimate to be one-third the cost of SCSI-based systems. A second iteration of the specification was launched, called Serial ATA II. This was announced in February 2002 and will allow vendors to further address the needs of networked storage segments via enclosure management and improvements in cabling and performance.

From its inception, SATA was designed to address a number of critical design limitations of ATA:

- Master/slave configuration limitation
- 2 drive-per-channel constraint
- Maximum bus performance 133 MB/sec with shared bus bandwidth
- 18" maximum cable length
- No growth road map

Designed to provide a "best of both worlds" solution, Serial ATA replaces the existing parallel bus with serial links, and adds features designed to increase the performance, reliability and scalability of ATA-based devices while retaining ATA's significant cost advantage. The result is a highly cost-effective alternative to SCSI technology for all but the most critical direct-attach and networked storage applications.

2. Gartner, 2002
3. HP Server Blades: Reducing Total Cost of Ownership, an IDC Whitepaper sponsored by Hewlett-Packard, November, 2001

3.0 Serial ATA Overview

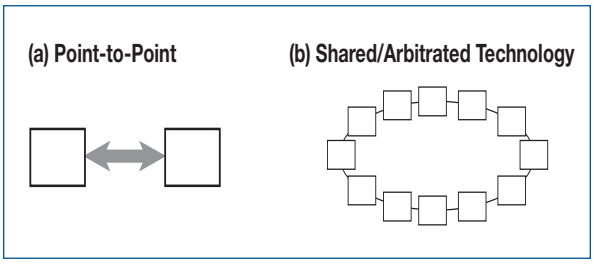
While SATA represents the next generation of ATA interface technology, it departs from ATA with several significant new features and benefits, summarized below.

Serial ATA Features (not in parallel ATA)	Benefits
Point-to-Point Configuration	Eliminated bus sharing overhead
Road map starts at 150MB/sec	Performance now and in the future
Additive device performance	Full bandwidth to each drive
1 meter cable length	Enables scalability
Hot-plug drives	Quick and easy drive replacement
Inexpensive disks	Lower system cost
Command queuing	Quick access to data
CRC	Strong data integrity

The following sections examine important new features of SATA, and how they can contribute to enhanced storage system performance, reliability and scalability.

3.1 Point-to-point Connectivity

This important connectivity feature provides significant performance and reliability advantages over the shared connectivity approach employed by both the ATA and SCSI parallel interfaces. Each port on a Serial ATA controller serves just one device; that is, the controller communicates with a given drive only through the port where it is connected (see figure below).



Because there is no sharing of the bus, each drive can communicate directly with the system at any time. As a result, the entire available interface bandwidth is dedicated to each device. This dedicated link approach eliminates the arbitration delays sometimes associated with shared bus topologies. With a shared bus approach, overhead increases as drives are added to the shared bus. This means that, in a typical ATA or SCSI RAID system, adding a disk will increase the total system throughput by some amount less than the throughput of the disk. With Serial ATA, on the other hand, each added disk can deliver its maximum throughput. Point-to-point connectivity offers the added benefit of simpler configuration. Dedicated links make a Serial ATA RAID system easy, fast and relatively inexpensive to set up.

Point-to-point connectivity introduces an important issue relative to scalability. With Serial ATA, the capacity supported by the configuration is a function of the number of available point-to-point connections. Traditionally, the architecture would determine the number of devices supported, and the number of connectors on the cable attached would dictate the number of drives that may be connected. With SATA, scaling is achieved simply by adding more point-to-point links in the system, at the host level, with each connector having one cable and one drive connected to it. The number of links included is dictated by the number a vendor has included on the system board. Additional links can be added via a controller card or a RAID card.

3.2 Cyclical Redundancy Checking (CRC) Error Detection

Perhaps the most beneficial and most significant improvement of Serial ATA over Parallel ATA is cyclical redundancy checking (CRC) covering each transfer. This feature, available in SCSI since Ultra160 technology was introduced, significantly improves data integrity.

With Serial ATA, each protocol layer has the capability to identify errors and can perform recovery and control actions as well as forward information to the next higher layer in the stack. Each layer has a means to be aware of – and recover from – errors in the layer below it. A detailed explanation of this feature is covered in section 4.1.

3.3 Hot-plug Support

SATA supports hot-plugging, the ability to swap out a failed disk drive without having to power down the system or reboot. This capability contributes to both data availability and serviceability, without any associated downtime. This feature is critical to extending SATA into enterprise applications.

The Serial ATA 1.0 specification requires staggered pins for both the hard disk drive and drive receptacles. Staggered pins mate the power signals in the appropriate sequences required for powering up the hot-plugged device. These pins are also specified to handle in excess of the maximum allowed inrush current that occurs during drive insertion. SATA-compliant devices thus need no further modification to be hot-pluggable and provide the necessary building blocks for a robust hot-plug solution, which typically includes:

- **Device detection** even with power downed receptacles (typical of server applications)
- **Pre-charging resistors** to passively limit inrush current during drive insertion
- **Hot-plug controllers** to actively limit inrush current during drive insertion

3.4 Improved Cabling

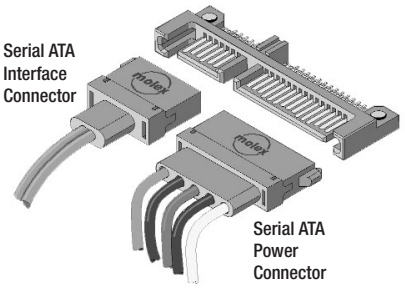
Cabling sits near the top of the list of hardware-related service calls. It also is a significant consideration in air flow design and hardware assembly. Serial ATA features an enhanced cable design that offers important benefits that add up to reduced service calls and efficient manufacturing.

Though it appears simple, the basic Serial ATA connector design is a remarkably efficient and practical design offering a number of notable features/benefits:

- The "L" shaped Serial ATA data and power connectors make plug orientation very obvious to the end-user, and prevent incorrect mating.
- The extrusion has "ears" which guide and align the plug during the mating process.
- The conductors are engineered for hot-plugging; they connect in three stages – first pre-charge, then ground, then power.
- The connector locations on the back of 2.5" devices are the same as for 3.5" devices, allowing design of backplanes that accommodate either size device.

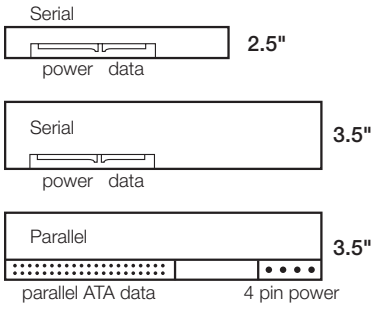
The Serial ATA connector represents a substantial improvement over Parallel ATA, which has a long history of problems with bent pins. And it represents a significant improvement over SCSI's daisy-chain topology, where, if one cable disconnects or fails or if a terminator is missing, the entire group of drives will not perform properly.

Appearance of Serial ATA Connectors



Drawing courtesy of Molex

Device Connector Sizes and Locations



4.0 Serial ATA Reliability

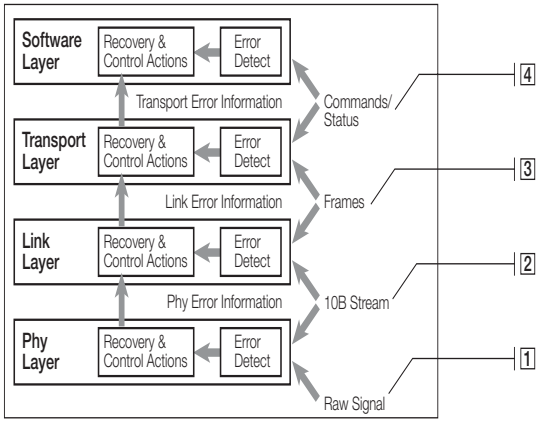
Serial ATA features greatly enhance storage system reliability as compared to ATA. Unlike its predecessor, SATA provides important protection and recovery features at three levels: the *interconnect* and *protocol* level, the *device* level, and the *system* and *campus* level.

4.1 Reliability at the Protocol and Interconnect Level

Thanks to its CRC feature (see section 3.2) SATA provides solid error detection at every layer:

- **PHY layer** detects and handles raw signal and 10-bit stream problem conditions such as no device, OOB signaling failures, and PHY internal errors. Information pertaining to these errors is also visible to the link system.
- **Link layer** detects and handles 10-bit stream and frame problem conditions such as invalid state and data integrity errors. Information pertaining to these errors is also visible to the transport layer.
- **Transport and software layers** detect and handle command and status problem conditions such as internal, frame, protocol, and state errors, including error handling for command block status registers and command failed/timeout. Information pertaining to the transport layer is visible to the software layer.

The diagram below helps to illustrate this feature:



As explained previously, SATA features hot-plug ability, plus an enhanced cable design that helps eliminate connection errors and bent pins. Together, these features contribute to improved reliability, reduced downtime and fewer services calls.

4.2 Reliability at the Device Level

It is reasonable to expect the reliability of SATA disk drives to approximate that of best-in-class ATA drives. A historical comparison of SCSI vs. Parallel ATA drive reliability, expressed as Mean Time Between Failure (MTBF), is shown in the table below.

Technology MTBF Comparison (Best in Class)

Year	SCSI MTBF	Parallel ATA MTBF
1997	1.0 million hours	500,000 hours
2002	1.2 million hours	800,000[4] hours

While SCSI drives clearly have a higher MTBF, it is useful to look a bit deeper at this metric. Consider that a MTBF of 800,000 hours (ATA) translates to a typical drive failing every 91 years. Also consider that a statistical mean provides information about a population, but not particular drives. The failure of a particular drive cannot be predicted. In other words, disk drive failure is an ever-present reality.

To address this reality, storage vendors utilize Redundant Array of Inexpensive Disks (RAID) to provide disk mirroring and parity data protection. In the 20 years since RAID was created, the "inexpensive" in the acronym has evolved to "independent," reflecting the evolution toward more expensive disks to achieve higher RAID performance.

Serial ATA offers price/performance that puts the "inexpensive" back in RAID. In addition to its low cost, Serial ATA has characteristics that make it particularly good for RAID, including dedicated point-to-point channels with a manageable cabling topology and staggered spin-up. The result is an affordable storage solution that delivers more than satisfactory uptime for a range of applications not considered "mission-critical."

4. Maxtor Corporation, DiamondMax Plus D740X.

ATA drives, with which Serial ATA drives will be price competitive, currently offer compelling products based on capacity and bandwidth at a lower cost than SCSI, as shown in the following:

Vendor	Maxtor	Western Digital	Seagate
Product Line	DiamondMax	WDCaviar	Barracuda
Model No.	4G120J6	WD2000BB	ST1181677LCV
Interface	EIDE UDMA 133	Ultra ATA/100	SCSI Ultra 160
RPM	5400	7200	10000
Capacity	160GB	200GB	180GB
Data Rate	133 MB/s	100 MB/s	160 MB/s
ASP	$250	$400	$1300

Source: www.nextag.com; www.maxtor.com; www.westerndigital.com; www.seagate.com

Likewise, with the introduction of 150 MB/s Serial ATA, a three drive point-to-point configuration of Serial ATA drives will provide 450MB/s of combined bandwidth compared to 320MB/s provided by the upcoming Ultra 320 SCSI drives at comparable costs.

4.3 Reliability at the Subsystem Level

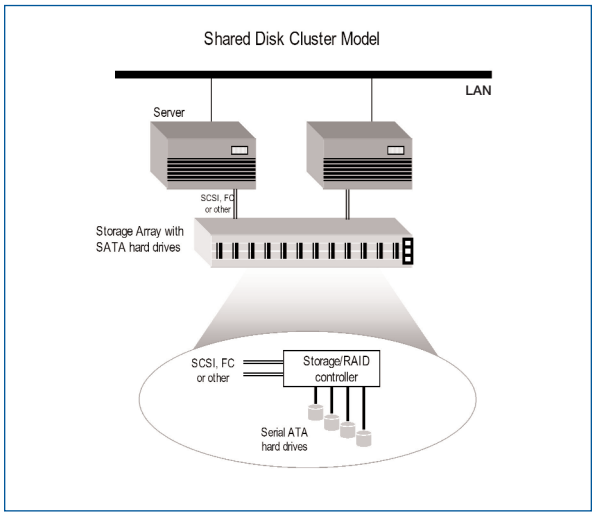
The physical hardware of a storage system also impacts its availability. The enterprise storage market is typified by storage arrays that provide dense, rack-optimized solutions with redundant components such as power supplies and fans. Systems manufacturers are striving for ever-denser solutions, requiring smaller form factors, to meet customer demand for more storage in limited space. This means the number of drives requiring cooling is increasing. This growing challenge for storage enclosure manufacturers is exaggerated when utilizing ultra-high-density SATA drives.

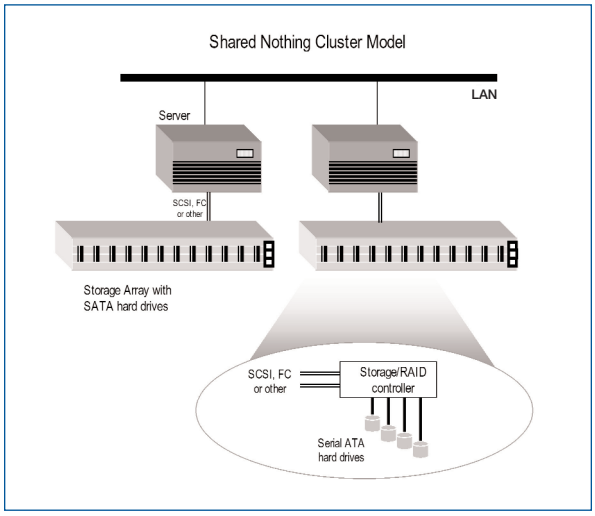
Temperature is a critical factor impacting the reliability of drives and other components. Simply speaking: the cooler the drives, the greater the reliability. As enclosure manufacturers develop solutions based on these ultra-high-density drives, it is critical to ensure proper cooling and fan redundancy to preserve system availability. SATA is well-positioned to handle these environments; the thinner cabling allows better airflow through the system.

4.4 Reliability at the System Level

To achieve high system availability (as well as scalability), many solution providers employ cluster configurations. There are two basic cluster configuration models: the *shared disk model* and the *shared nothing model*. Serial ATA can be employed effectively in either of these two models.

Both clustering models eliminate the server (or appliance) as single points of failure. And in both cases, Serial ATA RAID arrays can eliminate the disk drive as single point of failure.





4.5 Summarizing SATA Reliability

Based strictly on MTBF specifications, SCSI drives clearly deliver a higher level of uptime than SATA drives. However, it is system uptime that really matters. Since no hard drive is infallible, the only way to meet the need for 24 x 7 availability is to create a system that can tolerate a failure of any component, without loss of connectivity. The only way to achieve this is to create a system that has redundant *everything* – from multiple I/O controllers and servers to multi-path cabling.

Dual port SCSI drives can add value to such a redundant solution; if it is in a RAID volume, no data or connectivity is lost if the drive itself fails. However, a single port drive can achieve the same objective through RAID mirroring, such as RAID 5+1 (this is striping with parity + mirroring). Although mirroring requires twice as many drives, many IT professionals intentionally implement this method for enhanced performance, offering the ability to access the same data in parallel from both drives (~2x performance).

Additional technologies are also used to aid IT in the ability to predict failures and to quickly swap out faulty components. System solutions include such implementations as rigid backup policies, hot spares, hot-plug, enclosure management/error prediction (SMART), or any combination of these system solutions and others as determined by each business needs.

Thus, if one can achieve the system goal of guaranteed data reliability and data availability with either SCSI or SATA technology in a RAID array, then the deciding factor will likely be cost. In such a scenario, SATA offers an economic advantage that is hard to ignore.

5.0 Serial ATA Performance

Since devices based on SATA are still in development, the performance of these devices has yet to be fully tested. However, it is clear that SATA will usher in an entirely new level of ATA performance, beginning at 1.5 Gb/second (150 MB/second). This surpasses today's mainstream.[5] ATA performance by 50% or more, with a clear roadmap for future performance increases. Recall that in a system combining multiple devices, such as a RAID configuration, SATA offers the advantage of additive device performance, devoting full bandwidth to each drive in the system.

A typical misconception about SATA is that its serial nature makes it inherently slower than parallel bus designs. However, witness the trend toward serial technology for all desktop/server data transport mechanisms including PCI Express, Fiber Channel SANs, Infiniband Architecture, and others. Indeed, even SCSI is moving toward serial technology, with the recently introduced Serial Attached SCSI (SAS) draft standard. Given the high speeds required by today's technology coupled with the synchronization constraints of parallel data buses, high-speed serial links are the future.

5.1 How SATA Improves Performance

The key to SATA's higher performance (at least compared to ATA) is its point-to-point topology. SATA does not have to share the ATA bus as in the traditional ATA master/slave topology. Add to that SATA's dedicated 1.5 Gb/second (150 MB/second) maximum performance per device, and it can be seen that the bus already has room to spare when today's best-of-breed drives are hard pressed to deliver 100 MB/second. Note that this is just the starting point for Serial ATA; the specification developers are planning increased speed transitions for the technology over the next several years.

5. The most common ATA hard disk drives shipping today have a maximum throughput of 100 MB/sec. Only one drive vendor is known to be shipping a faster drive, at 133 MB/sec.

5.2 Evaluating Serial ATA Performance

While it is likely that first-generation SATA drives may not match the throughput of the best SCSI drives, it is important to place this in the context of real-world storage solutions. Few storage solution vendors actually use best-of-breed SCSI devices in their systems, due to their high cost. The SATA cost advantage enables vendors to use best-of-breed SATA products at a cost lower than if they used "mainstream" SCSI devices. This may in reality deliver performance comparable to or less than the best SATA units.

The important point about SATA is that it narrows the gap between ATA and SCSI performance, while retaining the traditional ATA price advantage. And in today's IT marketplace, for a growing number of storage applications, price/performance is the name of the game.

6.0 Conclusion

Clearly, SATA represents an important extension in proven ATA interface technology-one that targets SATA squarely at networked storage. The question is not whether SATA will penetrate the enterprise, but how far can it go?

While its attractive cost has already driven ATA technology into the enterprise space, SATA delivers performance and reliability that promise to solidify ATA's presence in the entry level/blade server category and extend to other business critical systems. SCSI will likely continue to be used for mission-critical storage applications. SATA, however, is ideally positioned to satisfy the requirements for a large number of enterprise storage applications, including:

- **Web hosting**
- **Firewalls**
- **Email servers**
- **Small business and Remote office locations**
- **File sharing**
- **Media streaming**
- **Tape backup replacement**

From a system point of view, SCSI and Fiber Channel will likely continue their hold on the high-end Storage Area Network (SAN) segment. However, despite the growth of SANs, Direct Attach Storage (DAS) still dominates the storage market. These DAS configurations are attractive to organizations seeking moderate performance at a reasonable cost. This is a market space where Serial ATA-based storage solutions can provide a superior value proposition, appealing strongly to what is still the largest segment of the external storage market.

DAS customers can take advantage of new and attractive options that Serial ATA offers. Now they can choose enterprise-class storage subsystems with enhanced power, packaging and cooling characteristics within a fully redundant/fault-tolerant configuration; a feature set superior to existing SCSI JBOD arrays. In addition, the use of low-cost Serial ATA external RAID controllers within such subsystems in combination with Serial ATA drives will provide a new entry point in terms of cost per Gigabyte in the RAID storage market space.

There will always be a need for storage solutions with higher performance and reliability than that delivered by ATA technology. The question for storage solution providers and customers is this:

Is the information of such critical importance that it warrants paying the significant premium SCSI- and Fibre Channel-based systems will command over a SATA based solution?

Given today's emphasis on IT costs, the exploding growth of storage capacity and the attractive performance roadmap and reliability features offered by Serial ATA, it is likely that many IT managers, especially in the Small and Medium Enterprise (SME) space, will say "Yes" to SATA for a growing number of enterprise storage applications.